Zurich Open Repository and Archive
University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2021

The Swiss StudentLife Study: Investigating the emergence of an undergraduate community through dynamic, multidimensional social network data

Vorös, Andràs ; Boda, Zsófia ; Elmer, Timon ; Hoffman, Marion ; Mepham, Kieran ; Raabe, Isabel J

Abstract: The Swiss StudentLife Study (SSL Study) is a longitudinal social network data collection conducted in three undergraduate student cohorts (N1 = 226, N2 = 261, N3 = 660) in 2016−2019. The main goal of the study was to understand the emergence of informal student communities and their effects on different individual outcomes, such as well-being, motivation, and academic success. To this end, multiple dimensions of social ties were assessed, combining computer-based surveys, social sensors, social media data, and field experiments. The dynamics of these social networks were measured on various time scales. In this paper, we present the design and data collection strategy of the SSL Study. We discuss practical challenges and solutions related to the data collection in four areas that were key to the success of our project: study design, research ethics, communication, and population definition.

DOI: https://doi.org/10.1016/j.socnet.2020.11.006

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-206581
Journal Article
Published Version

The following work is licensed under a Creative Commons: Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.

Originally published at:
Vorös, Andràs; Boda, Zsófia; Elmer, Timon; Hoffman, Marion; Mepham, Kieran; Raabe, Isabel J (2021). The Swiss StudentLife Study: Investigating the emergence of an undergraduate community through dynamic, multidimensional social network data. Social Networks, 65:71-84.
DOI: https://doi.org/10.1016/j.socnet.2020.11.006

Social Networks 65 (2021) 71–84

Contents lists available at ScienceDirect

Social Networks

journal homepage: www.elsevier.com/locate/socnet

The Swiss StudentLife Study: Investigating the emergence of an undergraduate community through dynamic, multidimensional social network data*

András Vörös [a,b,*], Zsófia Boda [a,c], Timon Elmer [a,d], Marion Hoffman [a], Kieran Mepham [a], Isabel J. Raabe [e], Christoph Stadtfeld [a]

[a] Social Networks Lab, ETH Zürich, Switzerland
[b] Department of Social Statistics and Mitchell Centre for Social Network Analysis, University of Manchester, United Kingdom
[c] Institute of Education, University College London, United Kingdom
[d] Department of Psychometrics and Statistics, University of Groningen, The Netherlands
[e] Institute of Sociology, University of Zürich, Switzerland

ARTICLE INFO

Keywords:
Dynamic networks
Multidimensional networks
Education
Network data collection

ABSTRACT

The Swiss StudentLife Study (SSL Study) is a longitudinal social network data collection conducted in three undergraduate student cohorts ($N_1 = 226$, $N_2 = 261$, $N_3 = 660$) in 2016–2019. The main goal of the study was to understand the emergence of informal student communities and their effects on different individual outcomes, such as well-being, motivation, and academic success. To this end, multiple dimensions of social ties were assessed, combining computer-based surveys, social sensors, social media data, and field experiments. The dynamics of these social networks were measured on various time scales. In this paper, we present the design and data collection strategy of the SSL Study. We discuss practical challenges and solutions related to the data collection in four areas that were key to the success of our project: study design, research ethics, communication, and population definition.

1. Introduction

Social relations influence individual outcomes in virtually all segments of society, including workplaces (Venkataramani et al., 2013), corporate boards (Westphal and Milton, 2000), entrepreneurial markets (Burt, 2009), healthcare institutions (Cunningham et al., 2012), rural and urban neighborhoods (Beggs et al., 1996), political groups (Johnson and Orbach, 2002), judiciary groups (Lazega et al., 2006), and criminal gangs (Morselli, 2010). The effects of peer relations on students in education have particularly attracted the attention of researchers in the past decades (e.g. Coleman, 1961; Sacerdote, 2011; Veenstra and Dijkstra, 2011). Peer social networks have been shown to matter for students' health behaviors (Simpkins et al., 2013; Haas and Schaefer, 2014), delinquency (Sijtsema and Lindenberg, 2018; Gallupe et al., 2019), empathy (Wölfer et al., 2012), career decisions (Anelli and Peri, 2017; Raabe et al., 2019), and academic achievement (Gremmen et al., 2019; Stadtfeld et al., 2019).

Collecting detailed and high-quality social network data is crucial to develop our understanding of how peer relations evolve and affect individual outcomes, such as the success and well-being of students. However, as social processes at school are highly complex, network data collection remains a major challenge for educational research to date. In particular, the fact that social networks are *multidimensional* and *dynamic* presents difficulties for empirical studies and data collection (e.g. Magnani and Wasserman, 2017; Lazega and Snijders, 2015). First, peer relations in the classroom or school are complex and take many different forms. Students may have friends and rivals, contacts within and outside of their class or cohort, and/or strong and weak relationships. The different types of social ties may be interrelated (Vörös and Snijders, 2017). For instance, students may perceive their friends as smart and

* All authors contributed equally to the paper. ZsB, IR, CS and AV wrote Sections 1, 4 and 5. TE, MH and KM wrote Sections 2 and 3 and created the figures and tables. TE and KM managed the Supplementary Information document. The authors planned the paper and evaluated progress together. All authors reviewed multiple drafts. AV coordinated the writing process.

* Corresponding author at: Department of Social Statistics and Mitchell Centre for Social Network Analysis, University of Manchester, United Kingdom
E-mail addresses: andras.voros@manchester.ac.uk (A. Vörös), christoph.stadtfeld@ethz.ch (C. Stadtfeld).

https://doi.org/10.1016/j.socnet.2020.11.006

Available online 15 December 2020
0378-8733/© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

funny, but less so those they dislike. Joint activities and time spent together strengthen the ties between individuals, but belonging to different groups may weaken them.

Second, these multidimensional peer relations change over time. This may be due to external factors, such as the arrival of a new classmate, a change in the curriculum, or developmental processes (exogenous change), or due to the structure of peer social networks themselves (endogenous change; Snijders et al., 2010). Moreover, the different dimensions of social ties might not change at the same pace. While close friendships or group perceptions may take months or years to develop, who one studies or spends free time with may change in the course of a few weeks or days (Kitts, 2014). In summary, studies that aim to understand the development and effects of social networks in schools need to collect data on multiple types of peer relations over time, and they should consider that important processes may occur on quite different time scales.

A number of data collection methods have been applied in educational research and beyond to tackle the issues of multidimensionality and dynamics in network data, which have specific strengths and weaknesses. *Surveys and interviews* are perhaps the most popular techniques to gather information on peer social relations and educational outcomes (e.g. Coleman, 1961; Snijders and Baerveldt, 2003). Relying on self-reports of social interactions and relations, these methods allow the large-scale collection of multidimensional and longitudinal network data, such as in the Add Health (Harris, 2013), the CILS4EU (the German and Swedish samples; Kalter et al., 2013), the KiVa (in Finland: Sentse et al., 2014; in The Netherlands: Rambaran et al., 2019), the PROSPER (Osgood et al., 2013), the SNARE (Laninga-Wijnen et al., 2019), and the RECENS (Boda and Néray, 2015) studies. However, due to cognitive limits and recall biases (e.g. Bernard et al., 1982; Marsden, 1990), the types of relationships surveyed and the maximal frequency of data collection rounds are rather low. Especially for questions regarding weekly, daily, or even hourly interactions, regular surveys are impractical.

The *experience sampling method* is more suitable to measure short-term dynamics with surveys or interviews, for instance, by asking students about important interactions of the past day at random times or every evening (Reis and Wheeler, 1991). In turn, the scope of these short surveys are necessarily limited and they cannot be repeated over long periods, such as months or years. Answering the same questions over a long time may lead to response fatigue and high rates of non-response.

A more recent approach to circumvent the limitations of regular survey designs employs digital technologies in an attempt to collect more "objective" data. S*ocial sensor techniques* rely on, for example, Bluetooth via smartphones (Sekara et al., 2016), Radio Frequency Identification (RFID) technology (Cattuto et al., 2010), or other sensor technologies such as infrared signals (Pentland and Heibeck, 2008) to collect fine-grained observational data on collocation and interactions between students. Social sensors do not need the attention of the participants, thus they can attenuate response fatigue and provide acurate data with high temporal resolution. Although data from social sensors and self-reports are positively correlated, sensor measures convey unique information that is otherwise difficult or impossible to gather (Eagle et al., 2009; Elmer et al., 2019). The direct observation of individual behavior allows to overcome measurement biases associated with self-reports (Bernard et al., 1982) and refocus research attention on social action (Baumeister et al., 2007). On the downside, social sensors are not suitable for measuring multiple types of social ties, only interactions based on physical proximity, and their use in studies with long data-collection periods may be technically difficult or infeasible due to limitations related to e.g. battery life, correct and regular use by participants, and malfunctions (although see Sekara et al., 2016).

Data from *online social media platforms* can fill these gaps, as they may provide real-time information on multiple types of social relations and interactions, even over longer time periods. Studies of online social networks have shown that these measures meaningfully complement measures of offline relations, and may provide, for example, different views on ethnic and gender segregation (Wimmer and Lewis, 2010; Hofstra et al., 2017). However, such contexts only record digital communication, ignoring offline social ties, while the two can only jointly provide a complete picture of interpersonal relations. Further, the algorithmic design and sampling of these platforms is often obscure, limiting data interpretability (e.g. Morstatter et al., 2013). Finally, as social media activity is highly sensitive information, it may be challenging for researchers to get participants, data providers, and ethics committees on board with gathering such data.

Finally, *field experiments* and *network interventions* have been used in combination with some of the strategies described above in order to better understand the effects of social networks on student behavior and outcomes (Valente, 2012; Hamm et al., 2014). While these interventions may pertain to multidimensionality and network dynamics, their effects should be studied on various time scales. As we have pointed out, this is difficult using data that come from only one of the above data collection strategies. Besides that, successful experiments and interventions often require the support of not only participants but institutional stakeholders as well, which may be challenging to acquire.

While each of the network data collection methods discussed previously has specific strengths, we argue that none of these techniques provide a complete picture of social network processes in educational settings on their own. Yet, their combination seems to be a promising way to uncover processes in multidimensional networks occurring at multiple time scales. This is precisely the strategy we followed in the Swiss StudentLife Study (SSL Study). Previous network studies have combined some data collection techniques in educational settings, such as surveys, phone call records, and social sensors (Eagle et al., 2009; Sekara et al., 2016) or surveys and social media sources (Hofstra et al., 2017). Our work contributes to this field by exploring a study design in which the strength of each data collection method helps to overcome the limitations of the others.

In this article, we present the research design and data collection strategy of the SSL Study. The study was conducted in three cohorts of engineering undergraduate students ($N_1 = 226$, $N_2 = 261$, $N_3 = 660$) at a Swiss university in 2016−2019. The data collection aimed at providing insights into the multidimensional and dynamic nature of social networks and their effects on the well-being, motivation, and academic success of students. The emergence and development of student communities were followed from the very first day they met throughout their first academic years. As a key strength of our study, social networks and student outcomes were measured on various dimensions and time scales, by the combination of regular network surveys, experience sampling techniques, social sensor technologies, social media data sources, and field experiments. The approach and research design can be useful to collect multidimensional and dynamic network data both in education and across various other social contexts. In the following, we provide an overview of the student cohorts, the research design and the measures, then we discuss practical challenges and solutions in four areas that were key to the success of our project: study design, research ethics, communication, and population definition.

2. Empirical context and samples

The Swiss StudentLife Study (SSL Study) is a high-resolution, multimethod longitudinal network data collection conducted by the Social Networks Lab at the ETH Zürich between 2016 and 2019. The aim of the project was to measure and understand the development of multidimensional social networks and their effects on the well-being, motivation, and academic outcomes of students in newly formed undergraduate communities. The study followed three cohorts of students in German-language engineering and natural science bachelor programs at a Swiss public university. Students are admitted to these programs without an entrance exam (a limited number of places are available for international students). However, they have to take comprehensive exams at the end of their first year to proceed with their studies. These exams are notoriously difficult: only about half of all students pass on their first attempt. Those who fail may either try again one year later or drop out of the university. The second and third years of the bachelor programs hold further challenging exams and practical courses. However, successfully finishing the three years promises access to prestigious master programs and job market opportunities. Due to the high-pressure environment and difficult examinations, students likely experience much stress for extended periods, which may negatively affect their psychological well-being and academic achievement. In turn, many are likely to develop ways of coping with these pressures, such as building strong friendships and finding appropriate social support, and experience much personal growth.

The three studied cohorts, which we label from here on as Cohort I, II and III, are from different departments and bachelor programs. Students in each cohort came from various parts of Switzerland and from abroad, and they rarely knew each other before the start of their first year. In all three cases, they quickly developed densely-knit social networks (see Fig. 4 for the case of Cohort I) as they had the majority of their classes together with their peers, relatively separated from other students at the university. Cohort I started their studies in September 2016, Cohort II and Cohort III in September 2017. Cohorts I and II were smaller, with 226 and 261 students respectively, and were further divided into four majors. Students in the same program had more shared courses but were in contact with the other peers in their cohort. Cohort III was larger than the other two, with 660 students, and had a single study program for all students.

For additional descriptive statistics of the cohorts, see Table 1. In general, our sample was predominantly male, with Cohorts I and II having a 60–65 % male proportion, and a greater male majority of approximately 85 % in Cohort III. Swiss nationals formed a large majority in all cohorts, with German nationals comprising approximately half of the remaining minority in all cohorts.

3. Data collection strategy

3.1. Overview

The dataset from the SSL Study covers three academic years for Cohort I, from September 2016 to September 2019, and two years for Cohorts II and III, from September 2017 to September 2019. It includes data from a series of long surveys, two series of intensive short surveys, a study on face-to-face interactions during a welcome weekend in Cohort I, social media sources, and two different field experiments. We particularly aimed at gathering fine-grained information during critical periods: the first semester and the weeks before the first-year exams.

We presented the aims and conditions of our study to students in all three cohorts at the introductory lecture on their first university day in September. After this, we invited them to register for computer lab sessions and answer the first long survey over the first three days (this is explained in detail in Section 3.2). These questionnaires are sequentially labeled “L[1+]”, independently for each cohort. We emailed participation codes to those who were not able to attend the lab sessions; with these, they could complete the survey on their own computer by the end of the first semester week. The long surveys were repeated every two months in the first semester and then every three to four months until the end of the third year (Cohort I) or second year (Cohorts II and III). For the later surveys, we distributed personalized links through email and text messages but did not organize further lab sessions.

Table 1
Basic descriptives for the three cohorts in the SSL Study.

	Total *N**	Starting *N***	Mean birth year (std. dev.)	% Female	% Non-Swiss	Number of nationalities
Cohort I	226	205	1996 (2.0)	37.8	22.7	29
Cohort II	261	207	1997 (1.6)	34.6	19.5	14
Cohort III	660	538	1997 (2.0)	12.3	12.1	40

Note. * All participants ever included in our records for the cohort. ** Number of participants in our records at the first long survey. All other descriptives are calculated for the total N.

The most intensive data collection was carried out in Cohort I, particularly during the first year. First, a field experiment was carried out to randomize initial meeting opportunities between students during two Student Introduction Days (SIDs) in June and July 2016 (Section 3.6). Second, in the first week of the program the university organized a welcome weekend for the new students. We invited those attending to wear RFID badges in order to observe their early interaction patterns (Section 3.4). Third, during the twelve weeks following the weekend, students received short questionnaires (sequentially labelled “S[1+]” or “A[1+]” depending on their content; Section 3.3) on their phones three times per week to report either their social or studying behavior. Fourth, we sent a second series of short surveys every week during the eight weeks preceding the final exams of the first year (sequentially labelled “SS[1+]”, Section 3.3).

Cohorts II and III participated in a field experiment at their SIDs in June and July 2017, as well as in a second experiment on the first day of university (Section 3.6). Apart from these, the second and third cohorts were only invited for the long surveys. Finally, we collected data on students’ networks on Facebook in the first cohort (Section 3.5). Fig. 1 presents the detailed timeline of the first study year in Cohort I. Fig. 2 provides an overview for all cohorts and the entire duration of the study. In Section 4.1, we discuss in detail key challenges and difficulties our team encountered during study design and management.

Additional student data, including enrollment, demographics, academic results and contact emails, were provided by the university under a strict confidentiality agreement. These data were regularly updated to ensure the accuracy of our population boundaries and survey invitation lists. Student data were securely stored and made accessible only to a subset of the research team. Every student received an anonymous identifier that was used to access their contact information for survey invitations and for creating linked, anonymized datasets of responses. We generated individual survey links in Qualtrics (survey platform) but distributed them with our own scripts to preserve the anonymity of participants. See Section 4.2 for a detailed discussion about the most important ethical and privacy challenges in our study.

We compensated students in Cohort I with 30 CHF (~30 USD) for each long questionnaire and 5 CHF (~5 USD) for the short ones. For Cohorts II and III, the compensation for a long questionnaire was initially set to 25 CHF (~25 USD) until the eighth questionnaire (L8), when it was also increased to 30 CHF (~30 USD). Students could retrieve this compensation by scanning a barcode at the university’s cash desk. The rates were in line with scientific compensation schemes of the university, corresponding to the typical student assistant hourly salary. In addition, participants were given a number of points for answering each survey, weighting their chances in a tombola. The points added were increased linearly in each consecutive long survey, increasing the incentive later, when the participation was expected to be lower. We drew this tombola eight times and offered the winners prizes such as food and drink vouchers and a grand prize worth 1000 CHF (~1000 USD). Finally, we gave a short presentation to Cohort I after the first year, as an alternative to material incentivization, where we showed a few aggregated and anonymized network descriptives and findings. We felt that the event was motivating for participants, in line with earlier findings on self-information (e.g. Harari et al., 2017). However, due to the lack of resources to deliver a carefully anonymized but still informative presentation, we decided not to repeat the event for the other two cohorts. We opted for the use of various material

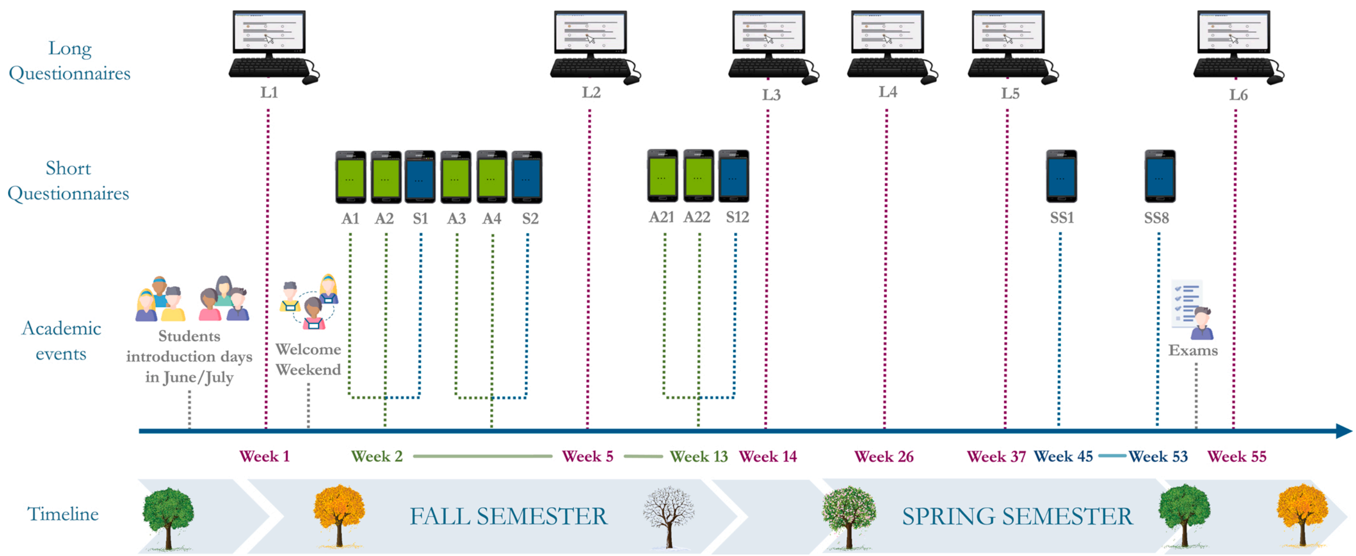


Fig. 1. Timeline of the first study year in Cohort I (September 2016 – September 2017). The courses of the Fall semester took place from Week 1 to Week 14 (mid-September to mid-December 2016) and of the Spring semester from Week 23 until Week 45 (mid-February to the end of June 2017). Weeks 45 to 53 (July and August 2017) were dedicated to the preparation of the final year exams in week 54 and 55 (end of August 2017).

incentives, because the size of the cohorts and our resources did not allow a strategy involving more personal contact, one we could follow in the RFID study and the field experiment(see Sections 3.4 and 3.6 for details), and because we wanted to minimize our intrusion on the social integration process.

Fig. 3 shows that with this compensation scheme, we maintained relatively stable absolute participation numbers in the first year of all cohorts, with some decline in later years. This was in part due to cohort changes as students dropped out of their study program, either for voluntary reasons or due to failing mandatory exams. Contrary to our expectation, cohort changes were difficult to track, even using official records, as students often notified the university only weeks or months after their decision to leave or join the program. We discuss the difficulties of population boundary definition in our study in Section 4.4. Table 2 reports indicative response rates for the long questionnaires based on our current best estimates.4.4Based on ad-hoc feedback from participants, the reasons for decreasing response rates may have been the lack of time (e.g. closeness of exams), the high frequency of surveys, and the fact that the surveys had quite similar content. We could address some of these issues, for example, by decreasing the frequency of surveys after the first study year and by introducing new question blocks (e.g. about cultural consumption and political opinions) in later long survey waves. In similar contexts, higher response rates could be achieved by varying the content of questionnaires and increasing the amount of informal communication with participants if possible. We discuss in Section 4.3 how we aimed to tackle challenges related to incentives and non-response with our communication strategy.

3.2. Long surveys

The core of our data collection was composed of repeated long surveys, with which we measured key individual and social network variables. The 45-minute online questionnaire contained up to 32 blocks of items. The surveys were administered using the Qualtrics software, through a server hosted and maintained by the Decision Science Laboratory (DeSciL) of the ETH Zürich. The survey items assessed seven broad aspects of the background and social life of participants: (1) demographics (e.g. gender, country of origin), (2) interpersonal networks (e.g. friendship, studying together), (3) mental health, identities, and personality (e.g. stress, identification with the university, Big Five traits), (4) attitudes towards politics (e.g. left-right self-placement), (5) free-time and cultural consumption behaviors (e.g. sports activities, substance use), (6) studying habits and attitudes (e.g. time spent on studying, motivation), and (7) social media use (e.g. frequency and duration of use of various platforms). Most items were included in all long surveys, but a small number of questions varied between waves and cohorts. Some of the varying questions were topical (e.g. asked about participation in specific university events), others were added were added in later waves based on the interests of new members of our research team. For a detailed summary of the blocks of questions, see Section S1 in the Supplementary Information (SI).

In each survey, participants had to agree to an informed consent text, only then they could start answering the questions. First, basic demographic characteristics were asked, such as participants' gender, country of origin and languages spoken with family/friends, if Swiss: region of origin, and financial status. Following these questions, participants reported on a predetermined set of social ties within the cohort (for a complete list including full text and translations, see Section S2 in the SI). These ties are conceptually grouped into four categories: activity ties, such as travelling or living together; perception ties, such as believing someone is likeable, arrogant, or smart; social role ties, such as who participants believed would be best suited to initiate social events, to de-escalate conflict, or organize the welcome weekend; and relational ties, such as being friends, supporting one another emotionally, or being in conflict. A hybrid name generator/roster approach was taken, with all current cohort-mates appearing in drop-down, autocomplete text fields. For each item, 5 or 20 alters could be nominated through these text fields (for details, see Section S2 in the SI). Fig. 4 shows the dynamics of the pleasant interaction and friendship networks between L1 and L3 in Cohort I (see Section S4 in the SI for plots with linked layouts). In addition, Fig. 5a presents the friendship network of L3 in the same cohort (using a different layout) for comparison with other network measures.

Further, we asked individuals about the informal social groups in the cohort that they felt they belonged to. This is similar to the approach of Socio Cognitive Maps (Cairns and Cairns, 1984; Neal and Neal, 2013). However, we focused on groups as perceived entities, and gathered more information about them than previous approaches. Participants named up to five groups, identified their members, and answered further questions about the groups, such as their typical social activities and individuals' self-identification with the groups. Fig. 5b shows the group nominations at L3 in Cohort I. A similar procedure was repeated for the

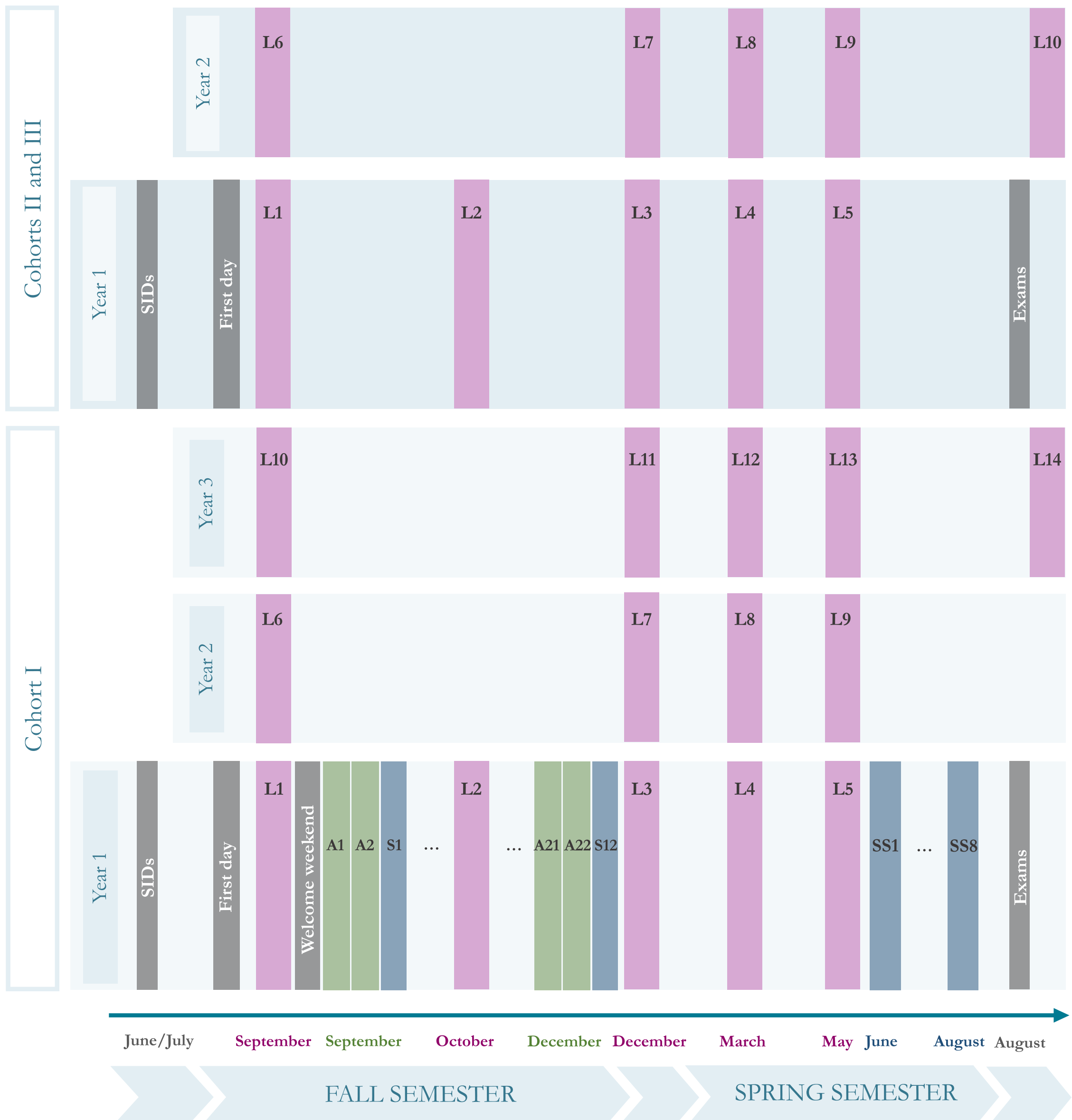


Fig. 2. Overview of the SSL Study in all three student cohorts. Year 1 is 2016 in Cohort I and 2017 in Cohorts II-III; SIDs = Student Introduction Days; First day = first semester day; L = long surveys; A = short surveys about affect and daily interactions; S = short surveys about weekly studying; SS = summer surveys about weekly studying.

perception of informal groups to which participants did not belong. Beyond these ties, we included the F-SozU scale (Kliem et al., 2015) in the surveys to assess perceived social support. Lastly, at two time points per cohort we asked participants about their most important relationships in the past half year, including those outside of their university community. We asked about the importance, roles, and attitudes these individuals held in the participants' lives. We did not ask about social ties between reported others ("alter-alter" ties), but relationships between reported university friends may be known if these friends were in our sample and responded to the network questions in the survey. For an illustration of these ego-centered networks, see Fig. 6.

After the collection of network data, we asked questions about participants' mental health. We focused on three aspects of mental health, anxiety, depression, and stress, as these were likely to be related to students' social integration and academic performance. We used validated scales and strictly followed standard protocols in psychology for collecting data about mental health (including giving feedback). In particular, the German versions of the GAD-7 for the assessment of anxiety (Spitzer et al., 2006; Löwe et al., 2008) and the CESD-R for depressive symptoms (ADS; Hautzinger and Bailer, 1993) were administered. Additionally, participants were asked whether they were in psychological treatment at the time and whether they wished to receive notification if their mental health was indicated to be poor. If they answered "yes" to the latter and scored above standard thresholds of the scales, we sent them an email within four weeks of the survey indicating our findings. We also administered the PSS-10 stress scale (Cohen and Williamson, 1988), and the Big Five Inventory in its 10-item (Rammstedt and John, 2007) and 44-item (Lang et al., 2001) forms.

Next, participants indicated their identification with several entities: their university, their department, the scientific community, their country of origin, men, women, their subject, and their canton. Identification was indicated by images of two increasingly overlapping circles (see Schubert and Otten, 2002), representing themselves and the entity in question.

A cultural consumption block followed this, in which we asked how frequently participants used substances, went to parties, and did sports.

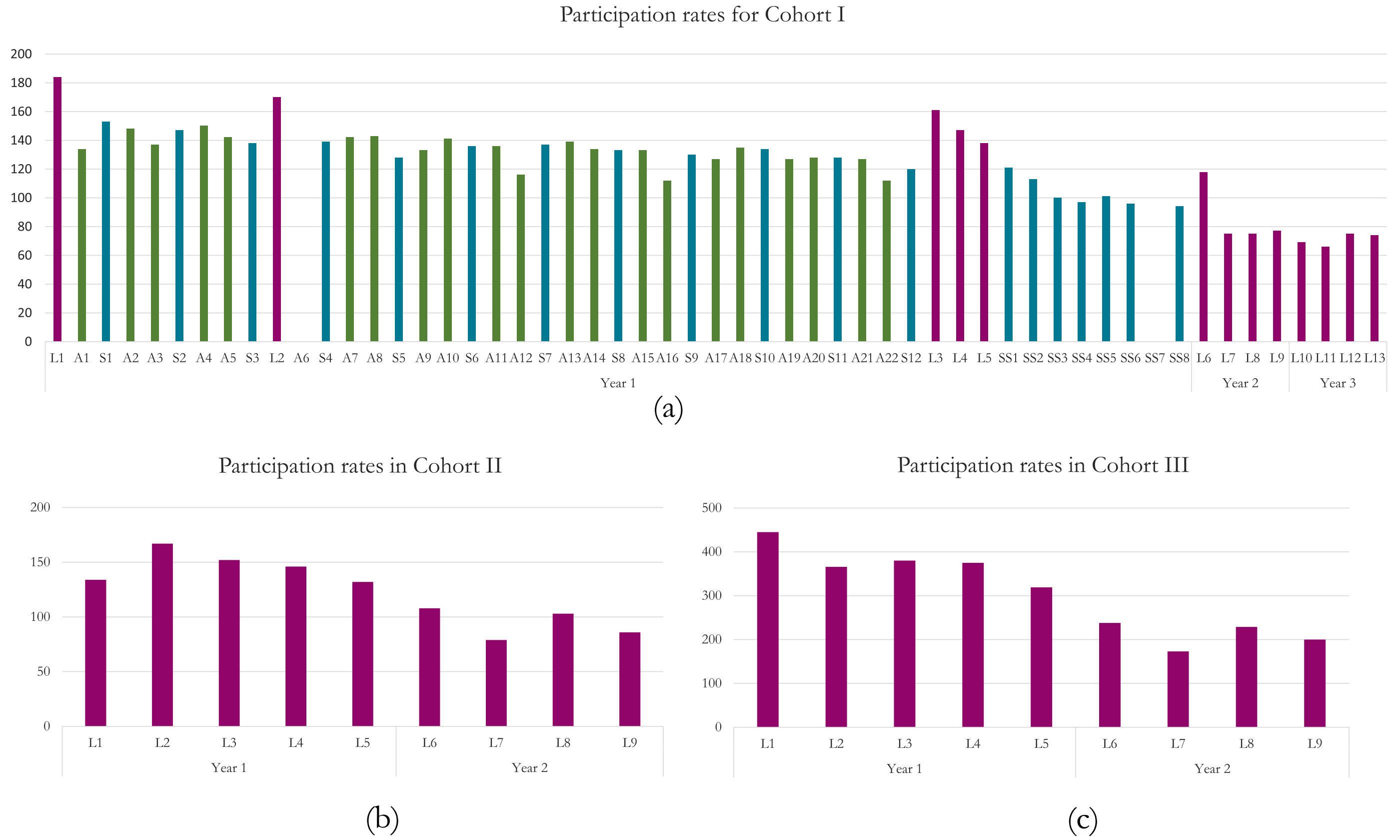


Fig. 3. Overview of participation in long and short surveys in the three cohorts. Two short questionnaires for Cohort I were not administered due to technical issues; starting and total population sizes are given in Table 1; see Section 4.4 for challenges in defining population boundaries in the study.

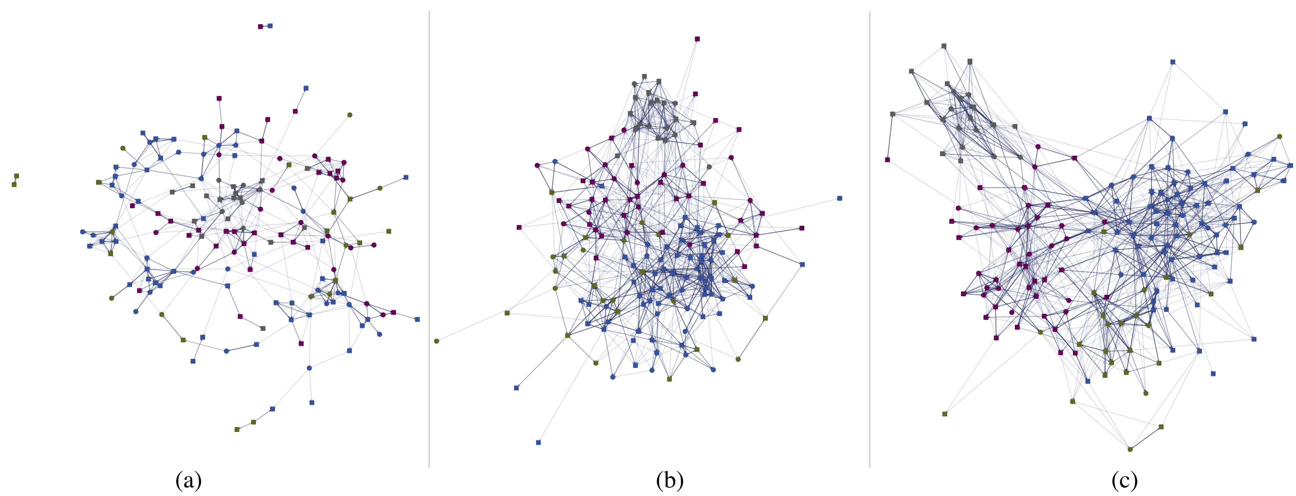


Fig. 4. The growth of the pleasant interaction and friendship networks over the first semester of Cohort I. Light grey edges are pleasant interactions, thick blue edges are friendships; data are from surveys (a) L1 (September 2016), (b) L2 (October 2016), and (c) L3 (December 2016); node color corresponds to the study program; square nodes are male, circle nodes are female; only individuals who participated or were nominated in all three surveys were included in the plots. *N = 171*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

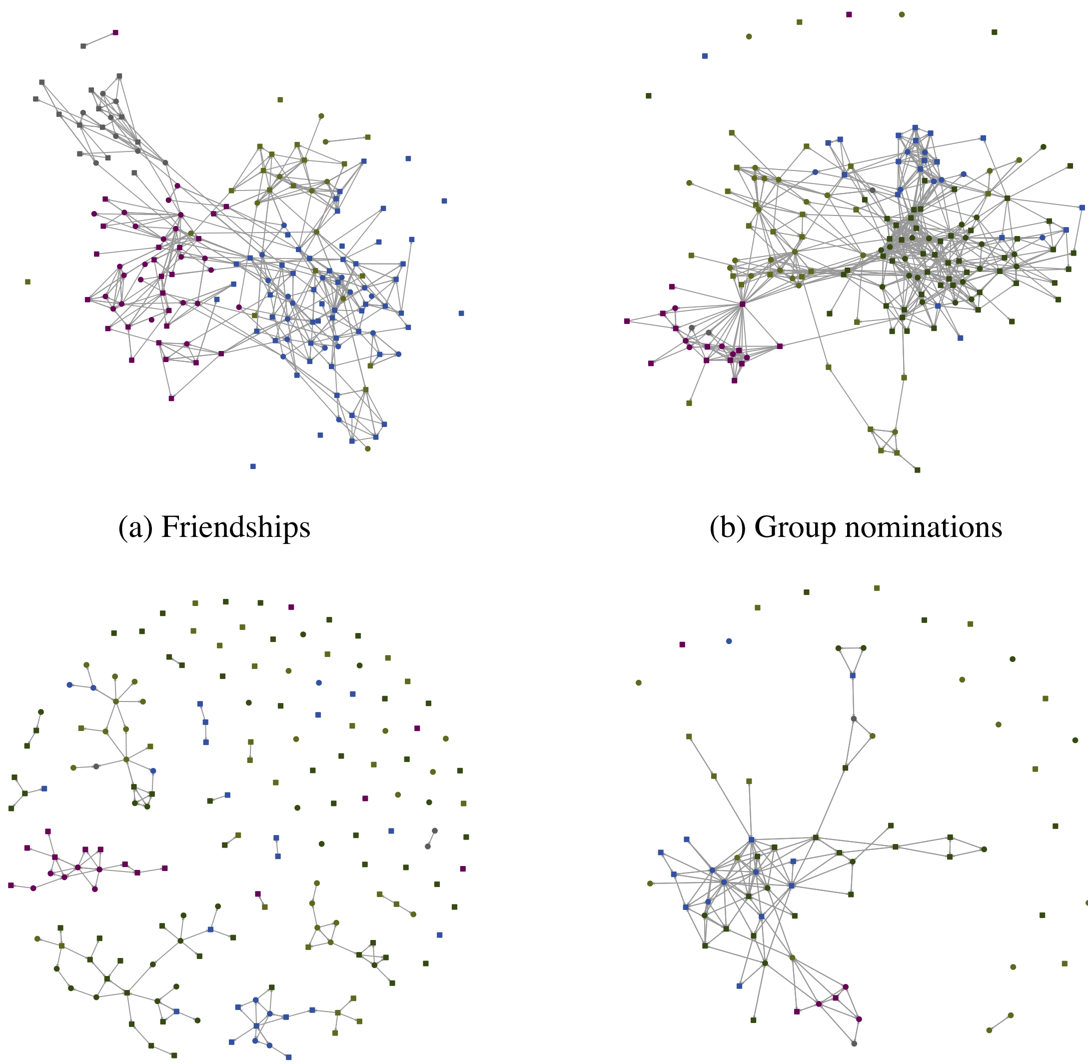
In Cohort I, we then asked them to indicate which music and movie genres they enjoyed most. In Cohorts II and III, we administered a modified version of the International Social Survey Programme's (ISSP Research Group, 2009) questionnaire on free-time expenditure and interest group membership. Some items in this block, especially those about substance use, could be considered highly sensitive by some of the respondents. We aimed to elicit responses by asking only a minimal amount of information, e.g. the frequency of cannabis use on a 6-point ordinal scale. This approach was successful: for example, averaging over all long survey waves, 94 %, 92 %, and 89 % of respondents answered the question about cannabis use in Cohort I-III, respectively.

Following this, questions were asked about participants' political orientations. Items included participants' self-placement on a left-right political scale, feeling thermometers regarding those from the left, right, and center, their political interest, their knowledge and agreement with topics subject to upcoming/past referenda, as well as their

Table 2
Estimated response rates for the long surveys.

	L1	L2	L3	L4	L5	L6*	L7	L8	L9	L10	L11	L12	L13
Cohort I	77 %	72 %	70 %	70 %	66 %	61 %	52 %	53 %	52 %	42 %	41 %	44 %	46 %
Cohort II	77 %	66 %	60 %	66 %	60 %	48 %	47 %	67 %	56 %	–	–	–	–
Cohort III	76 %	50 %	52 %	56 %	48 %	36 %	46 %	62 %	55 %	–	–	–	–

Note. Response rates are indicative, based on best estimates on population composition and change given available official records. See Section 4.4 for a discussion of related challenges. * Survey L6 was conducted in the weeks before exam results were announced and many students dropped out of their program, but they were officially still members of their cohort.



(a) Friendships

(b) Group nominations

(c) Short-survey interactions of one day

(d) Facebook friendships

Fig. 5. Four social networks at the time of survey L3 (December 2016) in Cohort I. The four networks are (a) friendship relations, (b) perceptions of being in the same informal group, (c) reports of daily social interactions from the 32nd short survey, and (d) Facebook friendships on December 19, 2016. $N = 171$. $N_{Facebook} = 74$.

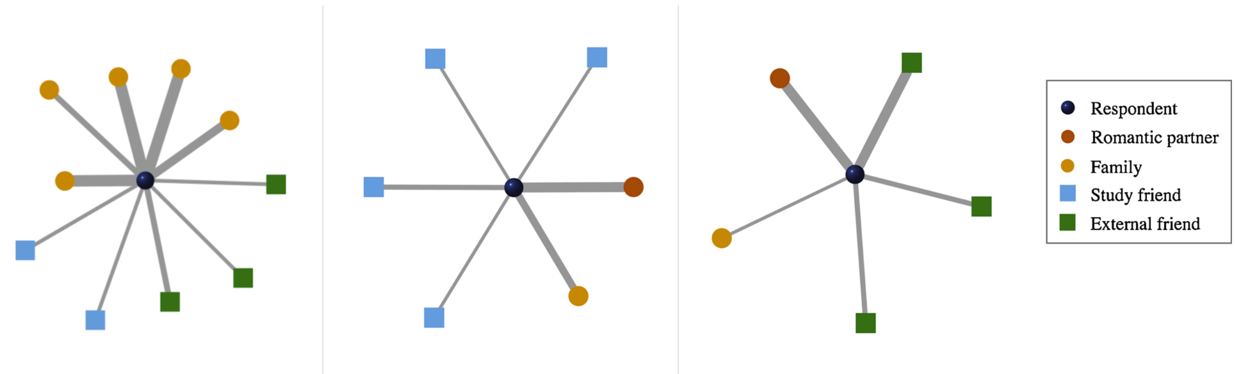


Fig. 6. Three example ego-centered networks from survey L4 (March 2017) in Cohort I. The respondent (ego) is marked by a dark blue circle; shapes and colors of other nodes (alters) represent relationship type: circles are family and partners, squares are friends (mutually exclusive categories); edge thickness indicates reported closeness to alters (on a scale from 1 to 7). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

intended/realized vote, and their political media consumption. Additionally, in Cohorts II and III, we presented participants with a battery of political statements sourced from the German Wahl-O-Mat (Wahl-O-Mat Research, 2017) and the schedule of upcoming Swiss referenda[1] and asked their agreement with the statements. Expecting that some of our respondents may find questions about political opinion highly sensitive, we used more detailed introductory notes in this block which reassured participants that there are no answers that are more correct or desirable than others. This approach appeared successful: for example, an average of 92 %, 95 %, and 92 % of respondents answered the question about political orientation in each wave in Cohort I-III, respectively.

We assessed participants' studying attitudes and behaviors in the next block, including their studying habits, intention to quit their degree, and their integration in relevant student associations, with an additional set of questions on perceived differences in scientific aptitude between men and women. Participants' work motivations were assessed with the German version of the SELLMO (Spinath et al., 2002; Wilbert, 2011).

An extensive block on social media use and perceptions was included in the later waves of the survey. This block included additional network questions on whose posts were seen, from whom they perceived the posts to be particularly good or bad, and with whom they communicated over various social media. The attitudinal and behavioral items asked about frequency and duration of participants' social media use, how important social media were for both their studies and social lives, and whether they perceived social media as having a net positive or negative influence on their lives. Participants in Cohorts II and III also completed the Emotion Recognition Index task (Scherer and Scherer, 2011). Finally, we asked participants for feedback on the survey.

Of Cohort I, 8 students (4%) completed none of the long surveys and 146 (65 %) completed more than 60 % of the 11 long surveys. In Cohort II, 8 students (3%) completed none of the long surveys and 174 (66 %) participated in more than 60 % of the 7 long surveys. In Cohort III, 22 students (3%) completed none of the long surveys and 310 (47 %) participated in more than 60 % of the 7 long surveys. Regarding feedback, across all cohorts and long surveys, the ratio of respondents finding the questionnaires enjoyable varied between 90 % and 40 % per survey, with a clear downward trend over time. Besides, between 50 % and 80 % found the questions rather similar in each wave. This highlights that repetitive question blocks were related to participant experience and, possibly, response rates. At the same time, the ratio of students who reported having considered to stop halfway through the questionnaire remained well below 40 % in all surveys. Finally, the amount of students reporting technical difficulties was consistently low, and almost all respondents found the material incentives fair in every wave. These figures are in line with the ad-hoc feedback we received from some of our participants, who found the survey generally interesting but noted the large number of questions and the similarity of survey waves. Overall, the feedback suggested that they found the surveys and incentives to be fairly aligned.

3.3. *Short surveys*

One of the primary aims of this data collection was to gather fine-grained information on how social interactions and (co-)studying occur on a daily and weekly basis. During two intensive observation periods that were particularly important for either the community-development (the first three months) or the studying process (the eight weeks before the first-year exams), participants responded to our short surveys. We implemented two different types of end-of-the-day surveys in Cohort I.

The first type of short survey was used only in the first semester. Participants received an invitation for this survey twice a week, at the end of varying weekdays. The survey asked students to report their five most important social interactions with other cohort members that day. For each of the interactions, we asked who was present, the start and end time, and the perceived pleasantness, energy, and intimacy of the interaction. These questions were inspired by the widely-used Rochester Interaction Record (Reis and Wheeler, 1991). Moreover, the students were asked about their affect (I-PANAS-SF; Thompson, 2007) and stress (PSS 4; Cohen et al., 1983). We varied the days of the week this survey was distributed over time in order to gain a more complete picture the daily interactions between students throughout the semester (and to avoid collecting data, for example, about the same study group meeting every week).

The second type of short survey was used in both intensive observation periods and it asked about various aspects of the participants' studying behavior. The surveys were administered at the end of each week (Sunday) of the first and the second intensive observation period. The participants were asked about how much they studied (on top of course attendance), how happy they were with this quantity, and with whom they studied.

Fig. 3 shows the participation rates of the two types of short surveys. Of Cohort I, 37 students (16 %) never participated in the short surveys and 128 (57 %) participated in more than 60 % of the 33 short surveys. Fig. 5c presents an example of the daily interactions reported by students in the 32nd short survey of the first semester (just before long

[1] There are four annual dates on which federal-level referenda votes are scheduled in Switzerland. These are typically about changes to the constitution which must always be voted on, laws proposed by the government which are disputed by >50,000 voters, or laws proposed by private individuals who are able to collect >100,000 signatures in favor of their change. This last kind is technically not a referendum, but a so-called popular initiative.

survey L3).

3.4. Social sensor (RFID) badges

During the first week of their program, a group of students from Cohort I participated in a welcome weekend organized by their student association. In total, this event involved 59 students and 14 additional students from the organizing team. To understand how these students interacted in this early stage of community development, we equipped them with RFID badges throughout the weekend. Such badges have been widely used to record face-to-face interactions and provide a reliable proxy for them in a relatively cheap, non-intrusive, and robust way (Cattuto et al., 2010; Elmer et al., 2019). They usually need to be worn on the chest and detect interactions by measuring if another badge is in close proximity and if the angle between the badges indicates that individuals are facing each other. More specifically, our badges recorded proximities up to 1.6 m and angles of about 65 degrees (Elmer et al., 2019). These records are collected by stationary routers through broadcast signals coming from the badges and the distribution of these routers defines the space in which interactions can be recorded.

Before the event, we installed this router system throughout the whole space used during the weekend and prepared an RFID badge for each participant. When arriving at the event venue, students were carefully informed about the badge's function and the purpose of the study. All of them agreed to wear one and were instructed to wear it for the duration of the whole weekend (from Friday at 7 pm to Sunday at 8 pm). We asked them to wear it on the chest to allow for optimal detection and allowed them to remove it when sleeping. The badges were covered with individual name tags and were therefore not visible. We also attended the event in order to ensure that participants were wearing their badge correctly and that the recording ran smoothly. In the beginning of the event, participants frequently talked about the RFID badges and made each other aware of its purpose of measuring social interactions. However, this initial excitement vanished quickly, within about two hours, and the participants did not seem very aware of the badges anymore. This is in line with research using video observations, which show that participants quickly forget that they are videotaped (Martin and Martin, 1984; Coleman, 2000). We regularly checked and found that participants did remember to wear their badges (and name tags) at all times.

It is important to assess the validity of RFID-based measures of social interactions before using the generated data to investigate substantive research questions. In a series of validation tests, Elmer et al. (2019) examine how good RFID badges are at capturing face-to-face social interactions, using the same hardware and technical setup as in our study. For example, pretests showed that hiding the sensors behind a thick paper or a plastic badge holder already reduces their range of detecting other badges dramatically. This suggests that a simple badge design with a thin paper name tag is the safest option – the one we opted for in the SSL Study. Further, battery life also appeared to matter: when batteries were more depleted, the range was reduced again. Therefore, we made sure that batteries were charged at the beginning of the data collection for the data not to be biased over the weekend. Elmer et al. (2019), furthermore, compared the RFID data to ground-truth video data from a lab study and self-reports of social interactions that were collected during the welcome weekend. It is also shown there how simple data processing strategies can improve the accuracy of the RFID badges.

3.5. Social media data

A significant part of a student's life involves being active on social media platforms such as Facebook. For this reason, we collected data about Facebook-friendship relations of Cohort I. Beyond this information directly from Facebook, an extensive block on social media use and attitudes was included in the long questionnaires. On Friday of the first week of university, all students of Cohort I were invited by email to participate in this part of the study. Participants were redirected to a website that informed them precisely what kind of data would be collected (i.e., only friendship ties to other students of the same cohort) and how their data would be treated by the research team. Then, participants were redirected to Facebook, where they would allow the StudentLife Facebook App to access their friendship data. Permissions were valid for 60 days, then they had to be renewed. For each permission request accepted, participants would be included in an extra tombola draw.

This StudentLife Facebook App granted us access to participants' friendship ties with other users of the application (i.e., other participants). Data on ties with individuals who were not part of the cohort or did not participate in this part of the data collection could not be collected. Facebook-friendship data of the participating students could be retrieved through the Facebook API with a valid access token that was created upon agreement to participate. Every morning at 2am our software would query the Facebook API with each valid access token for the friendship ties of participants, providing us with a daily update on the state of their network on the site (other data, such as relationship history were not retrieved). The collected data were then stored on a secure university-owned server.

If a student were not to renew their permissions, friendship ties to other "active" members would still be reported by the Facebook API. Therefore, there are technically three statuses of participation: (1) active (i.e., agreement to provide Facebook data at a given time point), (2) passive (i.e., once active but the access token was not renewed), and (3) non-participation (i.e., individuals that never agreed to provide access to their Facebook-friendship data). Fig. 5d shows the Facebook-friendship network of December 19th 2016 (i.e., during the collection of L3) for the $N = 74$ subset of students that participated in this part of the study.

The Facebook App was straightforward to implement based on official Facebook guidelines. The collection of Facebook data was subject to the same ethics guidelines we implemented in other parts of the study. Almost all students used their real names as a Facebook username, thus it was mostly relatively easy to link their Facebook data with their study ID. In the rare cases of unidentifiable Facebook usernames, the email address, that was part of the Facebook profile information, gave clear indications about the participants' real name. Once merged with the study database, identifying information in the data acquired by the Facebook App was deleted.

For Cohort II and III it was also planned to collect data from Facebook and additionally Instagram, but Facebook Inc. did not allow non-commercial data collection via APIs anymore at that time. At the time of writing this article, commercial companies are allowed to collect user data on Facebook and Instagram, whereas scientists are not (even though they pursue public interest and are obliged to follow strict ethical principles). We hope that these practices will be reviewed and modified in due course to promote broad data access and replicability in science.

3.6. Field experiments

In the scope of the study, two types of field experiments were carried out at the beginning of the first academic year of each cohort to understand the effects of early meeting opportunities on the emergence of social structure in student communities. First, students were randomly assigned to campus tour groups at Student Introduction Days (SIDs) organized by their university. This treatment was applied to Cohorts I and II. Second, students were randomly seated at their first lecture on their first semester day. This experiment was conducted in Cohorts II and III.

3.6.1. Student introduction days (SIDs)

Students in Cohorts I and II were invited to participate in one of two day-long information events organized for each cohort separately by the university. These took place 2–3 months before the start of the first academic year. The SIDs aimed to provide students with essential information about their upcoming studies and to host their first informal meetings with their new peers. Participation at the SIDs was voluntary and it did not represent an official commitment from students to register for the first semester.

The events began with an introductory lecture for all attending students. After this, participants were distributed into small groups for a tour of the university campus, a discussion and a meal led by an appointed tutor (a senior student). The groups consisted of 5–11 students. We had the opportunity to randomly assign participants to groups at the SIDs. One condition was given by the organizers: only students from the same study program could be assigned to the each group, as the induction was specific to the program. Tutors received the list of students who were assigned to their groups before the event, and they gathered their group after the introductory lecture. The number of students participating at the SIDs in Cohorts I and II respectively were 99 and 79.

While we implemented a conditionally random grouping, its results were imperfect. First, participation at the event was voluntary and hence not random (e.g. Swiss students were more likely to attend than students from abroad). Second, students could choose their preferred day of participation of the two days offered by the university. Third, not everybody who participated at the SIDs actually started their studies in the cohort 2–3 months later, as attendance at the event did not require an official registration for the first semester. Nonetheless, the data from the introduction days is still free of direct effects from implicit and explicit social preferences of students since they did not choose their groups.

3.6.2. *Seating at the first lecture*

In Cohorts II and III, we were able to randomize the seating of students during the introductory lecture on the first day of their studies. This was realized by a ticket system and hosts, akin to how seating works at theater or cinema. Each seat in the lecture hall was numbered by a sticker. As students arrived for the event, the researchers and trained assistants greeted them and gave them a badge containing a random seat number. Other members of the research team made sure that students found their seat and did not swap badges with others.

At the beginning of the lecture, students were asked to write their full name on the stickers at their seats. We took note of the names and seat numbers after the event, thereby recording the data on the randomized seating arrangement. The lecture was followed by a small-group tour similar to that at the SIDs. Participants were organized into groups for the tour based on the color of their randomly distributed badges. This provides us further information about which peers students spent their first hours with at university.

4. Challenges and solutions

In this section, we present some of the major challenges our research team encountered during the SSL Study. First, related to study design and management, we discuss how we pursued an ambitious and complex data collection with a small and interdisciplinary research team. Second, we describe our efforts to promote participation, data processing and data accessibility while maintaining the highest ethical standards and protecting the privacy of our respondents. Third, we present our communication strategy with participants and institutional stakeholders, which aimed at establishing and maintaining their trust and cooperation. Lastly, we highlight difficulties in identifying exact population boundaries, even in the highly organized setting of a university. We selected these specific challenges because we find that they are either not widely discussed in the literature or that their solutions require compromises between different research priorities.

4.1. *Study design and management*

4.1.1. *Study complexity vs. research team size*

The SSL Study has a complex structure consisting of multiple data collection methods. Certain study periods were especially intense in terms of effort and activity, such as the first semester of each cohort (in the Fall-Winter of 2016 and 2017). To tackle this problem, we found it useful to always assign responsible team members to each project part as well as each substantive topic during study design. The distribution of topics naturally followed from the research interests of the team members. During the execution of the study, we created a number of positions in our team:

- Project coordination (1 member): managing and balancing the different tasks and topics;
- Survey content management (1 member): coordinating topic leaders, compiling and updating surveys;
- IT and back-end management (1–2 members): handling data sets and online questionnaires;
- Data management (2–3 members): processing data, creating a codebook, handling changes in the sample;
- Communication (2 members): keeping in touch with study participants as well as stakeholders (university management, departmental administration, student organizations – see Section 4.3 below for further details about our communication efforts);
- Finances (1 member): overseeing the budget and expenses, such as assistant salaries and participant incentives.

Even though we constantly monitored the limits of the research team, there were peak periods (especially during the first weeks of the data collection) when the project was particularly stressful for the team members. On the one hand, the workload was higher than usual, on the other hand it was clear that potential errors would be more critical than in later phases of the data collection. In such periods it was important to reduce additional workload as much as possible. Researchers and PIs should pay attention to limits presented by team size and intense data collection periods in order to promote both project success and the well-being of their team members.

4.1.2. *Study length vs. changing team composition*

The initial planning of the study setup took place over an entire year in 2015–2016, with a team of 5 full-time researchers (1 assistant professor, 2 postdocs and 2 PhD students) and 4 student assistants. In the summer of 2018, at the start of the last and least intense data collection year, we had 6 full-time researchers (1 assistant professor, 2 postdocs and 3 PhD students) and 2 student assistants. While the research team did not change a lot in size, 7 changes occurred in-between (4 initial members left, 3 new members joined the team). The changes in team composition required shifts in roles and responsibilities. Ensuring the proper hand-over of tasks and training of new members took extra effort. While a long observation period is scientifically exciting, it comes with higher “turnover costs”. Research teams should reserve sufficient time and effort to accommodate changes in team composition.

4.1.3. *Diversity of research interests vs. sample fatigue*

The study was conducted by an interdisciplinary team with a wide range of research interests and considering a variety of disciplinary standards across the social sciences. The collected data has been used in several PhD and postdoc projects. Consequently, a large amount of relevant network and individual information had to be collected. However, this was obviously demanding for respondents, risking that they would opt out or provide poor-quality data. We took various steps to prevent sample fatigue. First, during the study-design phase we made sure that the research topic of each team member was sufficiently represented in the surveys. Securing the data necessary for successful PhD projects was a top priority. Decisions about measurement required detailed discussions and necessary compromises (e.g. excluding some survey items or using shorter scales where possible). Second, we aimed to evenly distribute the different data collections in time, especially the most demanding long surveys, short surveys, and the RFID data collection. We further avoided placing participants under unnecessary stress in intense studying periods, for example, by minimizing the length and complexity of short surveys in the weeks before their first-year exam. Third, using similar measures across different data collection methods did not only ensure the comparability of responses, but it also reduced the burden on participants. For instance, we employed similarly structured items to measure social networks in long and short surveys. Ultimately, we always gave participants a chance to easily drop out of the study or join again if they wished so. These steps resulted in a positive participant experience overall, based on feedback to the long questionnaires. Considering the long surveys across all three cohorts, 16 % (190) of students in our sampling frame participated in every survey after joining the study (though they might not have joined in the first wave), 38 % (443) joined at some point and dropped out later, 29 % (339) joined again after missing some surveys, and 17 % (202) never participated.

4.2. Research ethics and privacy

4.2.1. Network data quality vs. privacy protection

To collect high-quality network data, we needed to ensure that as many students as possible participated in our study and that alters nominated in network items could be unambiguously identified. At the same time, it was important to protect the privacy of our respondents. Reaching these goals is a difficult problem in most social network data collections, since respondents may report sensitive information about others who decided not to take part in the study (e.g. if they perceived them to be disliked in the community). This may lower trust in the data collection and lead to lower participation rates. We considered using free-text name generators (without suggestions from a pre-entered student list) for network items to eliminate some of these privacy concerns. However, such an approach in cohorts of several hundred students promised to lead to serious issues with identifying nominated alters, due to the use of inaccurate names or nicknames, and subsequently to a loss of network information. As a result, we opted for the hybrid name generator and roster method presented earlier, with dropdown autocomplete text fields. This strategy proved successful apart from a technical issue experienced by a few students in the first survey wave due to a browser setting. We could fix this problem by adding a short technical note to the questionnaire pointing respondents to the solution. Based on our experience, the hybrid method is easy to implement in the right survey tool and can be especially useful in larger cohorts. Smaller groups can be successfully surveyed using standard rosters.

To ensure broad participation, correct alter identification, and privacy protection with our chosen network data collection method, we developed a three-category participation scheme. At the beginning of the study and in each subsequent data collection step, students in our sample could choose whether they wanted to be active participants, passive participants or non-participants. They were asked to make this decision after being fully informed about the type of data they would be asked to provide. Active participants took part in the data collection (though they could naturally leave any survey question unanswered) and other participants could select them in network survey items. Passive participants did not take part in the data collection, but others could select them in network items. Non-participants did not take part and were not selectable by others in the online surveys; they were also removed from all subsequent communications about the study, upon request. Less than ten participants across the three cohorts chose this last option, either because they claimed to have dropped out of their studies, contrary to what official records suggested, or because they did not wish to receive our survey invitation and reminder emails. Overall, the scheme was received well by participants and was also accepted by our IRB as a solution to arising privacy concerns.

Introducing the passive participation option was crucial, as it allowed us to use the most up-to-date official student lists in our questionnaires, sometimes implemented only days before a given survey wave. For passive participation, we implemented a passive consent (opt-out) scheme, which required informing each participant in detail before each survey wave. In the case of Cohort I, students received information sheets about our study along with their official university admission letters. In the other cohorts, students were informed by e-mail. Team members attended the Student Introduction Days and the first-day welcome lectures where our PI presented the project. In the letters, presentations and questionnaires it was clearly communicated that students could opt out of passive participation at any time. Students could opt out through the official project e-mail or by contacting a team member directly. We were pleased that only a couple of students chose this option throughout the project, and mostly because they were not actively pursuing their studies. This suggests the success of the passive participation scheme and that our communication efforts were well-received in the cohorts.

For specific data collection steps, we implemented slightly different informed consent schemes. In the RFID data collection, students were informed upon arrival at the student weekend by a team member about the scope of the study. They had then the chance to review and sign an informed consent form. If they did so, they received a name tag with an RFID badge. If they had decided not to participate, they would have received exactly the same name tag (including the RFID badge) but with the battery removed in front of them. Without the battery, no data could have been generated. This protocol was never enacted, however, because all of the participants opted to take part in the RFID data collection. We believe this was the result of the physical presence of the research team at the weekend, who were able to openly and thoroughly explain the process and aims of the data collection to the participants. A few of our team members stayed at the event to set up and manage the RFID infrastructure. In the first few hours, they were frequently asked by participants about how the technology worked and how data was recorded and processed. They gave detailed answers and reminded the participants of our privacy and data security protocols. The same level of engagement would not have been possible if a larger group of students had attended the event or if the data collection spanned a longer time period. In the Facebook data collection, participants received an e-mail invitation to register through a Facebook app (see Section 3.5 for details). Only if they agreed to a consent form, would we collect data about their Facebook ties with other individuals of the cohort. Only information about pairs of participating individuals was recorded and stored by the app.

4.2.2. Efficient data processing vs. data security

Due to the large amount of data collected in different forms, it was important to ensure that the research team can work efficiently on the various dataset parts. This served to facilitate data processing, data cleaning and anonymization. At the same time, we aimed to store data from and about participants as securely as possible. Since it is necessary in longitudinal network studies to temporarily store identifying information about participants to link different surveys waves, the data were particularly sensitive. After studying the established good practices of social network research and consulting with our institutional ethics board, we established the following protocols. To ensure both efficient access and data security, we only worked on the non-anonymized data on secure, regularly backed-up university servers and never stored data using third-party storage (including Qualtrics). The most sensitive data was only accessible to a subset of the research team, who were in charge of managing the back-end and data processing. The dataset was anonymized by removing all personal identifiers and replacing them with randomized numerical IDs. For specific research studies, we only worked with subsets of the data that was necessary for the analyses. For limited feedback to participants and data sharing with other researchers, we only used anonymized and reduced versions of the dataset and reported only aggregated information.

4.3. Communication with participants and stakeholders

4.3.1. Collecting private information vs. establishing trust and cooperation

Communication and incentives were key to facilitate participation and good-quality data collection in our complex study. As we already introduced our compensation strategy in Section 3.1, here we focus on communication efforts. We had strict protocols in place for communication: the two team members charged with communication handed the email account set up for the study; only they communicated with participants, after discussing any issues with the rest of the team. We aimed to establish a personal connection with participants: all of our team members were present to meet them in person at the introductory events and the computer lab sessions of the first survey round, as well as at a feedback event organized for Cohort I after their first study year. The two researchers responsible for communication made personal visits to lectures to remind students of participation once before each long survey. They took a few minutes to present the project website and briefly explain the goals of the study, the participation options, the privacy policy, and the compensation scheme. To establish and maintain the trust of participants in the study, we further emphasized that we were working together with the university management, the relevant department, and student organizations, who were all informed of and approved our study. We reminded participants that these stakeholders do not have access to any data provided in the project.

We found that the above combination of communication efforts and incentives contributed to high participation rates and a positive overall participant experience (e.g. students could appreciate the purpose of the study, trust the professional conduct of the research team, and enjoy answering surveys). The importance of communication is perhaps best highlighted by the fact that we managed to achieve complete participation in the RFID study, where our research team was able to spend the most face-to-face time with participants and informally answer their questions about the project on the spot. Most of these questions pertained to the kind of data that can be collected with the RFID badges. We gave detailed responses and reminded participants of our anonymity and data security protocols. After a brief initial period of excitement, participants got used to wearing the badges and paid no mind to them during their interactions.

Beyond communication with the students, their participation was also supported by our regular contact with various stakeholders of the study, in both the design and the execution periods. To the management of the university, we gave frequent feedback and presented some results on various occasions. We worked closely with the administration who provided us with information about the study participants, including time-stamped enrollment lists and background data such as exam grades. We also received support from the relevant departments, especially by welcoming us at their events and helping us contact the students before their first semester.

We received ample support from departmental student organization, whom we approached in the study design phase. Our communications team presented them the aims of the study and answered their questions in an open meeting. They were especially interested in how our data and results could improve student experience in their program in the long run. Following the discussion, the organization members voted on whether they should endorse our study among the participants, which they overwhelmingly supported. Without this support, we would have likely experienced lower participation rates and we would not have been able to conduct the RFID data collection at the welcome weekend. Finally, we closely collaborated with the Decision Science Laboratory of ETH Zürich: we used their facilities and relied on their help and expertise in distributing surveys and managing compensation vouchers.

4.4. Population boundaries

4.4.1. Official vs. actual population composition

One of the largest challenges in the data collection phase of our study was the identification of population boundaries in our three cohorts. Even though we expected a quite orderly process of bachelor studies, it turned out that membership in study programs in higher education is not trivial. Many students enrolled in, but never started their studies. Who these people were only became known several weeks into the first semester. Quite a few other students joined later, some as late as year 2 or 3, as they were repeating a year after failing exams. Others were registered for multiple semesters at the same time. Finally, many students dropped out of their program permanently. However, they often announced this to the university at the beginning or end of a semester, and so the time between their decision (and actual withdrawal from the cohort) and their official sign-out is mostly unknown.

To tackle the issue of composition change and population boundary, we requested a large amount of meta-information from official university records, including exact enrollment and sign-out dates, course and exam taking, and so on. The analysis of these data is still in progress. For the moment, we record and report conservative estimates for participation rates as our working sample sizes are likely larger than the actual ones at any given time point (cf. Fig. 3 and Table 2). However, some uncertainty is likely to remain even after careful consideration of the detailed records, since official sign-up and drop-out dates might not perfectly reflect when a student joined or left the community (if at all). Therefore, in our research articles using the data, we also aim to explore other approaches to sample size estimation, for instance, taking into account that students who actually drop out are less likely to interact with their peers and should, therefore, be (nearly) isolated in many of our measured social networks. Our experience suggests that population boundaries may, in reality, be fuzzy and not necessary well represented by simple “ground truth” records. It seems to be a crucial practice to collect population information from as many sources as possible even in well-organized institutional settings.

5. Discussion

This paper presents the design and data collection strategies of the Swiss StudentLife Study (SSL Study), a high-resolution, multi-method longitudinal network data collection carried out in three bachelor cohorts of a Swiss university in 2016−2019. The main goal of the study was to measure the multidimensional and dynamic aspects of social networks between students and to understand how these networks impact individual outcomes. This was achieved by repeatedly assessing numerous types of social relations. Altogether, by asking a total of 1147 individuals, we collected information about 44 social network dimensions, and on dozens of individual characteristics, behaviors, and outcomes. Overall, we obtained long survey data at 14 (Cohort I) and 10 (Cohorts II and III) points in time, and short survey data at 40 points in time (Cohort I). To be able to measure social relations emerging in different periods and evolving at different paces, we deployed multiple data collection techniques: long online surveys, experience sampling-based short surveys, social sensor data, social media platforms, and field experiments.

By combining these strategies in the SSL Study, we could benefit from their respective strengths and collect detailed data on the emergence of peer communities and their effects on students’ lives. The potential of the dataset has already been explored in a number of research articles by our team, which focus on topics such as the effects of social integration on academic achievement (Stadtfeld et al., 2019), the effects of mental health on students’ social interactions at the welcome weekend (Elmer and Stadtfeld, 2020), the short- and long-term effects of early randomized meeting opportunities on friendship networks (Boda et al., 2020), and the validity of RFID sensor badges in measuring social interactions (Elmer et al., 2019). In addition, we complemented the dataset with two follow-up surveys in 2020 and studied impact of the COVID-19 pandemic on our students' mental health and social ties (Elmer et al., 2020).

Work currently in progress focuses on understanding the effects of randomized seating in the first lectures, the dynamics of friendship and political opinion, and the structure of group perceptions. Beyond these, the dataset allows the study of the effect of social integration and social influence processes on several types of academic outcomes (e.g. achievement, motivation, dropout) and indicators of well-being (e.g. stress, anxiety, depressive symptoms). For this, researchers can focus on the role of various aspects of social networks (e.g., friendships as well as support relations, negative ties, or those the person belongs to the same informal group with). Our data thus provides opportunities for a deeper understanding of the role of social ties within the student cohort on outcomes of its members.

We faced a number of challenges during our project and did our best to come up with appropriate solutions. We identified three main issue areas for which we could propose suitable solutions. First, the design and team infrastructure had to be optimized when the data collection was significantly extended. In this context, it was necessary to pay particular attention to team and project management. We found it very useful to include members in our team with experience in project management and social network data collection. Second, setting up the technical aspects of the data collection, the online surveys and the back-end, was a new task for our team. We were aware of the potential pit-falls, and managed to avoid major technical failures by continuously monitoring the systems involved in collecting and storing the data and by continuous data processing. Third, the necessity to gain and maintain the trust and support of the participants, student organizations, and university and department management was clear from the beginning. To that end, we developed a communication strategy that combined the open and repeated communication of study goals, prompt reactions to participant feedback, information events, monetary incentives and non-monetary prizes, as well as close contact with institutional stakeholders. Participation rates stayed satisfactory, especially in the two smaller cohorts.

We also faced some issues that remain partially unresolved. First, defining exact population boundaries appears difficult if at all possible. Second, while we had good overall participation, we had difficulties reaching some segments of the larger cohort. This demonstrates the potential need for even more personal contact with the students, and for developing alternative communication strategies. Third, while we did our best to specify research questions and identify the data necessary to answer these before starting the study, this turned out only to be possible to a certain extent, due to new members joining the team. We experienced a clear trade-off between flexibility in the survey content and optimal questionnaire design.

To put our data collection in context, our project belongs to a growing and diverse group of educational network studies. Well-known data sets such as the Add Health in the U.S. (Harris, 2013) or the CILS4EU in Europe (Kalter et al., 2013) commonly used by network researchers have the advantage of following large and heterogeneous cohorts. The SSL Study is much smaller in scope and represents a particular social context. At the same time, it includes very detailed and rich information on dimensions and dynamics of relationships in a few, relatively large, groups of students. We are convinced that various aspects of our study can be instructive for researchers planning future (large-scale) educational network studies. We would especially find a combination of large samples with detailed case studies carried out on their subsets a promising approach. This would allow the formulation and immediate large-scale testing of novel hypotheses about social network effects.

Our study highlights a number of potential avenues for the future development of network data collection in education. First, the measurement of multidimensional dynamic networks using the traditional survey approach is still often conducted in an ad-hoc manner; new standardized relational measures need to be constructed and tested. Second, high-resolution social interaction measures using surveys should be combined with the ever-developing smartphone-based technologies for the measurement of collocation; this may help to improve the validity and interpretability of both types of measures. Third, more work is necessary for the validation of different social sensor technologies; this work has to rely on the use of other measurement approaches. Fourth, the linking of social media data to interaction data gathered by experience sampling (such as in our short surveys) provides a promising way to understand the interrelations between short-term online and offline social behavior. Finally, the implications of field experiments have to be very carefully examined both in laboratory and real-life settings to establish and test causal links between social network structures and individual outcomes in education. We believe that the SSL Study presents another small step in the direction of complex social network studies, both in and beyond education, which can provide new insights into the role of social networks in our lives.

5.1. Data availability

The data of the SSL Study are available for replication and original studies. On the project website, we will publish the detailed codebook, scientific publications, replication material, and anonymized subsamples of the data. Due to privacy concerns, the full data are not available online. Access to sensitive variables can be granted on site and after a review of the research proposal. The data access policy is detailed on the project website. Its implementation is line with the open data policy of the Swiss National Science Foundation and the requirements of the institutional review board.

Acknowledgements

We thank Krishna Chaitanya, Charlotte Corrodi, Julia von Fellenberg, Prateek Purwar, and Afke Schouten for their hard work and contribution to the data collection. We are grateful to Stefan Wehrli (DeSciL at ETH Zürich) for his help in distributing our surveys. We also thank the university management, the department staff and the student organizations supporting the study and, above all, our participants for their enthusiasm, cooperation, and patience. The Swiss StudentLife Study was supported by the Swiss National Science Foundation(grant no. 10001A_169965,10DL17_183008 and P2EZP1_188022) and the rectorate of ETH Zürich.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.socnet.2020.11.006.

References

Anelli, M., Peri, G., 2017. The effects of high school peers' gender on college major, college performance and income. Econ. J. https://doi.org/10.1111/ecoj.12556.
Baumeister, R.F., Vohs, K.D., Funder, D.C., 2007. Psychology as the Science of Self-Reports and Finger Movements: Whatever Happened to Actual Behavior? Perspect. Psychol. Sci. 2 (4), 396–403.
Beggs, J.J., Haines, V.A., Hurlbert, J.S., 1996. Revisiting the rural-urban contrast: personal networks in nonmetropolitan and metropolitan settings. Rural Sociol. 61 (2), 306–325.
Bernard, H.R., Killworth, P.D., Sailer, L., 1982. Informant accuracy in social-network data V. An experimental attempt to predict actual communication from recall data. Soc. Sci. Res. 11 (1), 30–66.
Boda, Zsófia, Elmer, Timon, Vörös, András, Stadtfeld, Christoph, 2020. Short-term and long-term effects of a social network intervention on friendships among university students. Scientific Reports 10, 2889. https://doi.org/10.1038/s41598-020-59594-z.
Boda, Zs., Néray, B., 2015. Inter-ethnic friendship and negative ties in secondary school. Soc. Networks 43, 57–72.
Burt, R.S., 2009. Structural Holes: the Social Structure of Competition. Harvard University Press.
Cairns, R.B., Cairns, B.D., 1984. Predicting aggressive patterns in girls and boys: a developmental study. Aggress. Behav. 10 (3), 227–242.

Cattuto, C., Van den Broeck, W., Barrat, A., Colizza, V., Pinton, J.F., Vespignani, A., 2010. Dynamics of person-to-person interactions from distributed RFID sensor networks. PLoS One 5 (7), e11596.
Cohen, S., Williamson, G., 1988. Perceived stress in a probability sample of the United States. In: Spacapam, S., Oskamp, S. (Eds.), The Social Psychology of Health: Claremont Symposium on Applied Social Psychology. Sage, Newbury Park, CA.
Cohen, S., Kamarck, T., Mermelstein, R., 1983. A global measure of perceived stress. J. Health Soc. Behav. 24, 385–396.
Coleman, J.S., 1961. The Adolescent Society. Free Press, New York.
Coleman, T., 2000. Using video-recorded consultations for research in primary care: advantages and limitations. Fam. Pract. 17.
Cunningham, F.C., Ranmuthugala, G., Plumb, J., Georgiou, A., Westbrook, J.I., Braithwaite, J., 2012. Health professional networks as a vector for improving healthcare quality and safety: a systematic review. BMJ Qual. Saf. 21 (3), 239–249.
Eagle, N., Pentland, A.S., Lazer, D., 2009. Inferring friendship network structure by using mobile phone data. Proc. Natl. Acad. Sci. 106 (36), 15274–15278.
Elmer, T., Chaitanya, K., Purwar, P., Stadtfeld, C., 2019. The validity of RFID badges measuring face-to-face interactions. Behav. Res. Methods. https://doi.org/10.3758/s13428-018-1180-y. Advance online publication.
Elmer, Timon, Mepham, Kieran, Stadtfeld, Christoph, 2020. Students under lockdown: Comparisons of students' social networks and mental health before and during the COVID-19 crisis in Switzerland. Plos One 15 (7), e0236337.
Elmer, Timon, Stadtfeld, Christoph, 2020. Depressive symptoms are associated with social isolation in face-to-face interaction networks. Scientific Reports 10, 1444. https://doi.org/10.1038/s41598-020-58297-9.
Gallupe, O., McLevey, J., Brown, S., 2019. Selection and influence: a meta-analysis of the association between peer and personal offending. J. Quant. Criminol. 35 (2), 313–335.
Gremmen, M.C., Berger, C., Ryan, A.M., Steglich, C.E., Veenstra, R., Dijkstra, J.K., 2019. Adolescents' friendships, academic achievement, and risk behaviors: same-behavior and cross-behavior selection and influence processes. Child Dev. 90 (2), e192–e211.
Haas, S.A., Schaefer, D.R., 2014. With a little help from my friends? Asymmetrical social influence on adolescent smoking initiation and cessation. J. Health Soc. Behav. 55, 126–143. https://doi.org/10.1177/0022146514532817.
Hamm, J.V., Farmer, T.W., Lambert, K., Gravelle, M., 2014. Enhancing peer cultures of academic effort and achievement in early adolescence: promotive effects of the SEALS intervention. Dev. Psychol. 50 (1), 216–228.
Harari, G.M., Müller, S.R., Mishra, V., Wang, R., Campbell, A.T., Rentfrow, P.J., Gosling, S.D., 2017. An evaluation of students' interest in and compliance with self-tracking methods: recommendations for incentives based on three smartphone sensing studies. Soc. Psychol. Personal. Sci. 8 (5), 479–492.
Harris, K.M., 2013. The Add Health Study: Design and Accomplishments. Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill.
Hautzinger, M., Bailer, M., 1993. Allgemeine Depressions-Skala: Testmappe. Beltz.
Hofstra, B., Corten, R., Van Tubergen, F., Ellison, N.B., 2017. Sources of segregation in social networks: a novel approach using Facebook. Am. Sociol. Rev. 82 (3), 625–656.
ISSP Research Group, 2009. International social survey programme: leisure time and sports - ISSP 2007. GESIS Data Archive, Cologne. ZA4850 Basic Questionnaire. https://doi.org/10.4232/1.10079.
Johnson, J.C., Orbach, M.K., 2002. Perceiving the political landscape: ego biases in cognitive political networks. Soc. Networks 24 (3), 291–310.
Kalter, F., Heath, A.F., Hewstone, M., Jonsson, J.O., Kalmijn, M., Kogan, I., Van Tubergen, F., 2013. Children of immigrants longitudinal survey in Four European countries (CILS4EU). GESIS Data Archive, Cologne, ZA5353 Data File Version 1.0.0.
Kitts, J.A., 2014. Beyond networks in structural theories of exchange: promises from computational social science. Advances in Group Processes. Emerald Group Publishing Limited, pp. 263–298.
Kliem, S., Mößle, T., Rehbein, F., Hellmann, D.F., Zenger, M., Brähler, E., 2015. A brief form of the perceived Social Support Questionnaire (F-SozU) was developed, validated, and standardized. J. Clin. Epidemiol. 68 (5), 551–562. https://doi.org/10.1016/j.jclinepi.2014.11.003.
Lang, F.R., Lüdtke, O., Asendorpf, J.B., 2001. Testgüte und psychometrische Äquivalenz der deutschen Version des Big five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen. Diagnostica 47 (3), 111–121.
Laninga-Wijnen, L., Gremmen, M.C., Dijkstra, J.K., Veenstra, R., Vollebergh, W.A., Harakeh, Z., 2019. The role of academic status norms in friendship selection and influence processes related to academic achievement. Dev. Psychol. 55 (2), 337.
Lazega, E., Lemercier, C., Mounier, U., 2006. A spinning top model of formal organization and informal behavior: dynamics of advice networks among judges in a commercial court. Eur. Manag. Rev. 3 (2), 113–122.
Lazega, E., Snijders, T.A. (Eds.), 2015. Multilevel Network Analysis for the Social Sciences: Theory, Methods and Applications, Vol. 12. Springer.
Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., Herzberg, P.Y., 2008. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med. Care 46 (3), 266–274.
Magnani, M., Wasserman, S., 2017. Introduction to the special issue on multilayer networks. Netw. Sci. 5 (2), 141–143.
Marsden, P.V., 1990. Network data and measurement. Annu. Rev. Sociol. 16 (1), 435–463.
Martin, E., Martin, P.M.L., 1984. The reactions of patients to a video camera in the consulting room. J. R. Coll. Gen. Pract. 34 (268), 607–610.
Morselli, C., 2010. Assessing vulnerable and strategic positions in a criminal network. J. Contemp. Crim. Justice 26 (4), 382–392.
Morstatter, F., Pfeffer, J., Liu, H., Carley, K.M., 2013. Is the sample good enough? Comparing data from twitter's streaming API with twitter's firehose. Seventh International AAAI Conference on Weblogs and Social Media (ICWSM).
Neal, J.W., Neal, Z.P., 2013. The multiple meanings of peer groups in social cognitive mapping. Soc. Dev. 22 (3), 580–594.
Osgood, D.W., Feinberg, M.E., Gest, S.D., Moody, J., Ragan, D.T., Spoth, R., et al., 2013. Effects of PROSPER on the influence potential of prosocial versus antisocial youth in adolescent friendship networks. J. Adolesc. Health 53 (2), 174–179.
Pentland, A., Heibeck, T., 2008. Honest Signals. MIT Press.
Raabe, I.J., Boda, Z., Stadtfeld, C., 2019. The social pipeline: how friend influence and peer exposure widen the STEM gender gap. Sociol. Educ. 92 (2), 105–123.
Rambaran, J.A., Dijkstra, J.K., Veenstra, R., 2019. Bullying as a group process in childhood: a longitudinal social network analysis. Child Dev. https://doi.org/10.1111/cdev.13298.
Rammstedt, B., John, O.P., 2007. Measuring personality in one minute or less: a 10-item short version of the Big five Inventory in English and German. J. Res. Pers. 41 (1), 203–212.
Reis, H.T., Wheeler, L., 1991. Studying social interaction with the rochester interaction record. Adv. Exp. Soc. Psychol. 24 (C), 269–318. https://doi.org/10.1016/S0065-2601(08)60332-9.
Sacerdote, B., 2011. Peer effects in education: how might they work, how big are they and how much do we know thus far?. In: Handbook of the Economics of Education, Vol. 3 Elsevier, pp. 249–277.
Scherer, K.R., Scherer, U., 2011. Assessing the ability to recognize facial and vocal expressions of emotion: construction and validation of the Emotion Recognition Index. J. Nonverbal Behav. 35 (4), 305.
Schubert, T.W., Otten, S., 2002. Overlap of self, ingroup, and outgroup: pictorial measures of self-categorization. Self Identity 1 (4), 353–376.
Sekara, V., Stopczynski, A., Lehmann, S., 2016. Fundamental structures of dynamic social networks. Proc. Natl. Acad. Sci. U. S. A. 113 (36), 9977–9982.
Sentse, M., Kiuru, N., Veenstra, R., Salmivalli, C., 2014. A social network approach to the interplay between adolescents' bullying and likeability over time. J. Youth Adolesc. 43 (9), 1409–1420.
Sijtsema, J.J., Lindenberg, S.M., 2018. Peer influence in the development of adolescent antisocial behavior: advances from dynamic social network studies. Dev. Rev. 50, 140–154.
Simpkins, S.D., Schaefer, David R., Price, Chara D., Vest, A.E., 2013. Adolescent friendships, BMI, and physical activity: untangling selection and influence through longitudinal social network analysis. J. Res. Adolesc. 23, 537–549. https://doi.org/10.1111/j.1532-7795.2012.00836.x.
Snijders, T.A.B., Baerveldt, C., 2003. A multilevel network study of the effects of delinquent behavior on friendship evolution. J. Math. Sociol. 27 (2–3), 123–151. https://doi.org/10.1080/00222500305892.
Snijders, T.A., Van de Bunt, G.G., Steglich, C.E., 2010. Introduction to stochastic actor-based models for network dynamics. Soc. Networks 32 (1), 44–60.
Spinath, B., Stiensmeier-Pelster, J., Schöne, C., Dickhäuser, O., 2002. Skalen Zur Erfassung Der Lern-und Leistungsmotivation: SELLMO. Hogrefe.
Spitzer, R.L., Kroenke, K., Williams, J.B., Löwe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch. Intern. Med. 166 (10), 1092–1097.
Stadtfeld, C., Vörös, A., Elmer, T., Boda, Z., Raabe, I.J., 2019. Integration in emerging social networks explains academic failure and success. Proc. Natl. Acad. Sci. U. S. A. 116 (3), 792–797.
Thompson, E.R., 2007. Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). J. Cross. Psychol. 38 (227-), 242.
Valente, T.W., 2012. Network interventions. Science 337 (6090), 49–53.
Veenstra, R., Dijkstra, J.K., 2011. Transformations in adolescent Peer networks. In: Laursen, Brett, Andrew, Collins W. (Eds.), Relationship Pathways: From Adolescence to Young Adulthood. Sage, New York, pp. 135–154.
Venkataramani, V., Labianca, G.J., Grosser, T., 2013. Positive and negative workplace relationships, social satisfaction, and organizational attachment. J. Appl. Psychol. 98 (6), 1028.
Vörös, A., Snijders, T.A., 2017. Cluster analysis of multiplex networks: defining composite network measures. Soc. Networks 49, 93–112.
Wahl-O-Mat Research, 2017. Wahl-O-Mat Bundestagswahl 2017. Results of the Online Questionnaire. available at: www.phil-fak.uni-duesseldorf.de/wahl-o-mat/online-befragungen/.
Westphal, J.D., Milton, L.P., 2000. How experience and network ties affect the influence of demographic minorities on corporate boards. Adm. Sci. Q. 45 (2), 366–398.
Wilbert, J., 2011. Die Einsetzbarkeit der Skalen zur Erfassung der Lern-und Leistungsmotivation (SELLMO) bei Schülern des Förderschwerpunkts Lernen. Zeitschrift für Heilpädagogik 1, 11–17.
Wimmer, A., Lewis, K., 2010. Beyond and below racial homophily: ERG models of a friendship network documented on Facebook. Am. J. Sociol. 116 (2), 583–642.
Wölfer, R., Cortina, K.S., Baumert, J., 2012. Embeddedness and empathy: how the social network shapes adolescents' social understanding. J. Adolesc. 35, 1295–1305. https://doi.org/10.1016/j.adolescence.2012.04.015.